



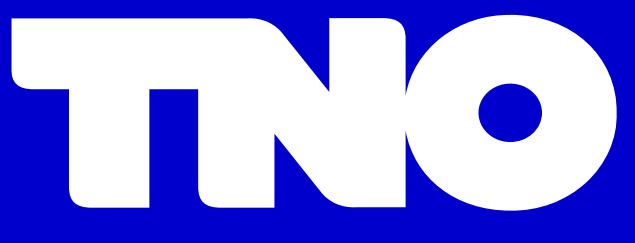
Proteomics an agnostic approach to detect viral species

Armand Paauw^{1*}, Manon Balvers², Bart D.J F Keijser³, Ingrid Voskamp-Visser¹, Jasper J. Kieboom³, Hans C. van Leeuwen¹, Evgeni Levin²

¹Netherlands Organization for Applied Scientific Research TNO, Department of CBRN Protection, Rijswijk, The Netherlands.

²HORAIZON Technology BV., Delft, The Netherlands

³Netherlands Organization for Applied Scientific Research TNO, Department of Microbiology and Systems Biology, Leiden



*E-mail: armand.paauw@tno.nl

Introduction

Identification of novel pandemic threats is a technological challenge. Efficient and rapid detection of novel pathogens and pathogen variants is of key importance in order to effectively institute counter measures and establish diagnostic protocols. In this project, we aimed to develop a proteomics-based workflow for efficient virus and strain variant identification.

Methods

Figure 1 describes schematic the developed pipeline, which exists of 3 steps.

1. Proteins in samples are purified and digested using single-pot, solid-phase-enhanced sample-preparation (SP3) technique (Hayoun *et al.* 2019).
2. The obtained peptides are then subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) on an Orbitrap instrument coupled to a nano-LC.
3. The obtained data is processed in PEAKS X to identify the peptides. To identify the peptides a proteomic database was constructed. The database contains the proteins of 1895 different virus species. The obtained peptide list is used to identify the virus(es) in the sample using the Proteome for pandemic preparedness (P4PP) application service on Proteome2Infection.com.

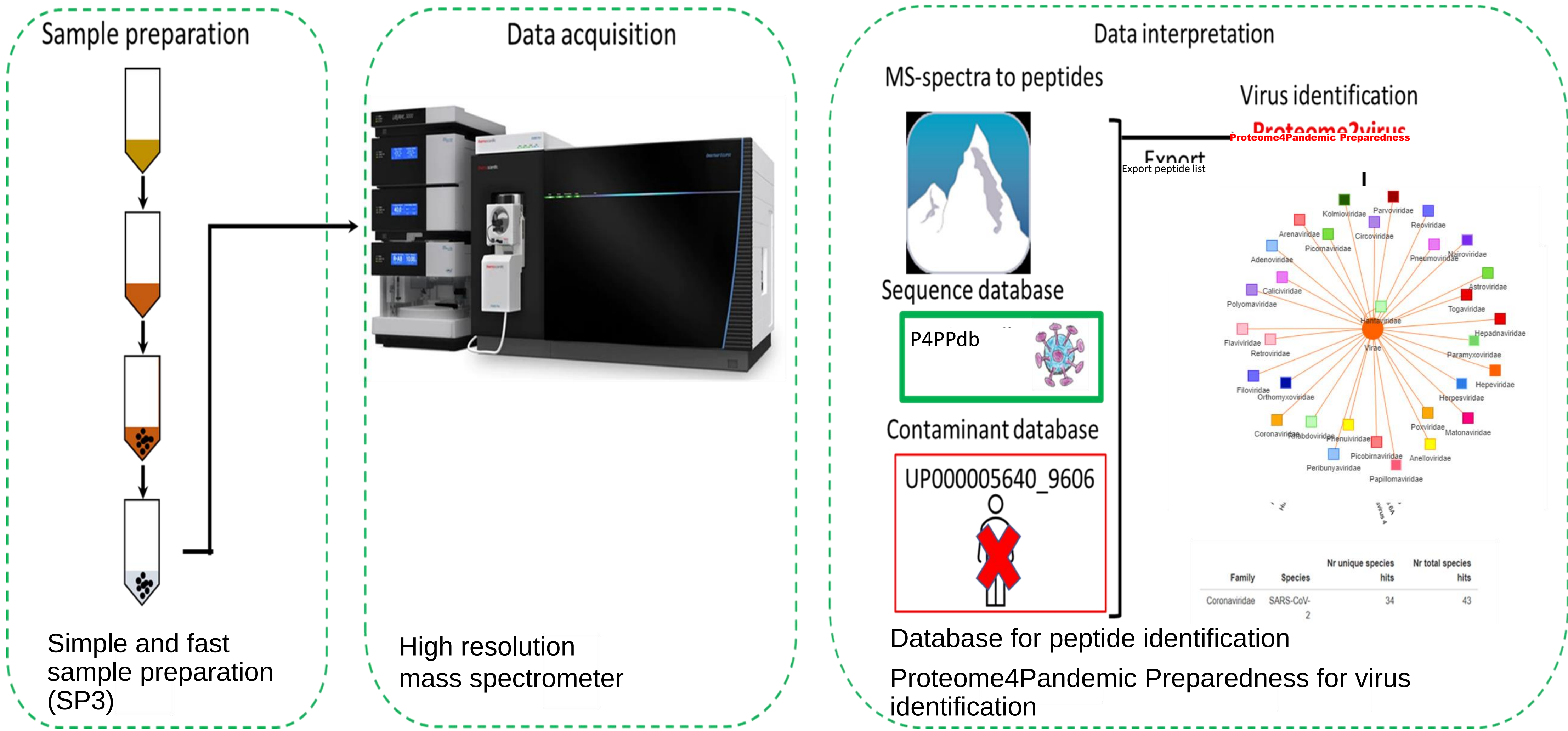


Figure 1. Schematic presentation of the proteomics pipeline for virus identification

Table 1. Results of Influenza A and SARS-CoV-2 samples cell cultures tested identified to species level and subtyped if possible.

Sample nr.	Identified virus	Based on X nr. of unique hits	Subtype	Based on X nr. of total hits
1	SARS	51	Alpha	49
2	SARS	37	Gamma	36
3	SARS	70	Delta	70
4	SARS	71	Beta or Mu	71
5	SARS	26	Mu or B.1.1.318	27
6	SARS	14	undetermined*	15
7	Influenza A virus	251	H1N1	241
8	Influenza A virus	45	H1N1	43
9	Influenza A virus	117	H3N2	115
10	Influenza A virus	17	H3N2	17
11	Influenza A virus	38	H3N2	38
12	Influenza A virus	46	H1N1	46

Acknowledgment

This work was supported by the Dutch MoD [grant number V2207] and by the Dutch Ministry of Economic Affairs [ERP Pandemic Preparedness].

Results

To test the P4PP application, heat inactivated virus samples, consisting of 36 samples derived by cell cultivation and 8 fecal clinical samples, were tested. These viruses were previously identified by PCR or were certified inactivated samples from Helvetica Health Care (Switzerland). The developed P4PP analysis starts with a family-level search followed by a species-level search in each relevant family from the first step. If Influenza A or SARS, is identified an additional search is executed to subtype the strain. Six SARS-CoV-2 and six Influenza A were identified to the species level. Both species can be subtyped if enough unique peptides are detected (Table 1). Other virus species samples were identified (Table 2) In total 42 samples were analyzed and all results were correct in comparison with provided information by the company or PCR results.

Table 2. Results of other viral cell cultures (^) of clinical fecal samples (*) tested identified to species level and two negative controls (").

Sample nr.	Identified virus	Based on X nr. of unique hits	Sample nr.	Identified virus	Based on X nr. of unique hits
13	Parechovirus A ^	5	29	Camelpox virus ^	3
14	Influenza B virus ^	203	30	West Nile virus ^	6
15	Rhinovirus A ^	5	31	Sindbis virus ^	27
16	Human orthorubulavirus 4 ^	59	32	Pixuna virus ^	21
17	Enterovirus A ^	14	33	Betacoronavirus 1 ^	90
18	Measles morbillivirus ^	221	34	Betacoronavirus 1 ^	81
19	Mamastrovirus 1*	5	35	Human coronavirus 229E ^	42
20	Norwalk virus*	5	36	Human coronavirus 229E ^	36
21	Mamastrovirus 1*	3	29	Human orthopneumovirus ^	73
22	Rotavirus A*	221	30	Human orthopneumovirus ^	54
23	Human mastadenovirus F*	76	37	Human alphaherpesvirus 1 ^	140
24	Rotavirus A*	7	38	Human alphaherpesvirus 3 ^	199
25	Norwalk virus*	20	39	Human metapneumovirus ^	81
26	Human mastadenovirus F*	93	40	Human metapneumovirus ^	18
27	Vaccinia virus ^	20	41	Negative control 1 "	-
28	Cowpox virus ^	4	42	Negative control 2 "	-

Conclusions

These results show that proteomics is a very effective method to detect and identify a wide variety of viruses. SARS-CoV-2 and Influenza A can be subtyped if enough unique peptides are detected (Table 1). Other viruses are identified to the species level based on at least 3 unique peptides (Table 2). No false positive results were reported.

Advantages of our proteomics to detect emerging viruses

- Screening for multiple viruses at once
- Simple (One method for all types of viruses: DNA or RNA / enveloped or non-enveloped)
- No specific reagents (primers or antibodies)
- Reliable (identification ≤ 3 identified unique peptide sequences)

